

APPENDIX B

BOIES SCHILLER FLEXNER LLP

David Boies (*pro hac vice*)
333 Main Street
Armonk, NY 10504
(914) 749-8200
dboies@bsfllp.com

**LIEFF CABRASER HEIMANN &
BERNSTEIN, LLP**

Rachel Geman (*pro hac vice*)
250 Hudson Street, 8th Floor
New York, New York 10013-1413
(212) 355-9500
rgeman@lchb.com

JOSEPH SAVERI LAW FIRM, LLP

Joseph R. Saveri (SBN 130064)
601 California Street, Suite 1505
San Francisco, California 94108
(415) 500-6800
jsaveri@saverilawfirm.com

*Counsel for Individual and Representative
Plaintiffs and the Proposed Class*

(Additional counsel on signature page)

**CAFFERTY CLOBES MERIWETHER &
SPRENGEL LLP**

Bryan L. Clobes (*pro hac vice*)
135 S. LaSalle Street, Suite 3210
Chicago, IL 60603
(312) 782-4880
bclobes@caffertyclobes.com

DICELLO LEVITT LLP

Amy Keller (*pro hac vice*)
10 North Dearborn Street, Sixth Floor
Chicago, Illinois 60602
(312) 214-7900
akeller@dicellolevitt.com

Matthew Butterick (SBN 250953)
1920 Hillhurst Avenue, #406
Los Angeles, CA 90027
(323) 968-2632
mb@buttericklaw.com

UNITED STATES DISTRICT COURT

NORTHERN DISTRICT OF CALIFORNIA

SAN FRANCISCO DIVISION

RICHARD KADREY, et al.,

Individual and Representative Plaintiffs,

v.

META PLATFORMS, INC.,

Defendant.

Case No. 3:23-cv-03417-VC

**THIRD AMENDED CONSOLIDATED
COMPLAINT**

Class Action

Demand for Jury Trial

Plaintiffs Richard Kadrey, Sarah Silverman, Christopher Golden, Ta-Nehisi Coates, Junot Díaz, Andrew Sean Greer, David Henry Hwang, Matthew Klam, Laura Lippman, Rachel Louise Snyder, Jacqueline Woodson, and, Lysa TerKeurst, and Christopher Farnsworth (“Plaintiffs”), on behalf of themselves and all others similarly situated, bring this Class Action Complaint (the “Complaint”) against Defendant Meta Platforms, Inc. (“Meta”).

I. OVERVIEW

1. “Llama” is the name given to a series of large language models ~~created and maintained by Meta~~ (or “LLMs” for short) created and maintained by Meta. Meta has poured billions of dollars into developing its Llama models because these models are critical to the company’s commercial aspirations. By releasing its models open-source, Meta aimed to build a quick and vast userbase that would give it a competitive edge in an AI market in which it was lagging. Beginning with Llama 2, Meta has incorporated Llama into its full suite of commercial products, now including Facebook, WhatsApp, Instagram, and Ray-Ban’s Meta Glasses. Meta believes that its multi-billion dollar investment in generative AI technology through its Llama models will bolster and define the company’s commercial and competitive advantage in the market for years to come.

~~1.2.~~ A large language model is an AI software program designed to emit convincingly naturalistic text outputs in response to user prompts. In this Complaint, unless otherwise indicated, “Llama” refers to all versions of Llama in any stage of their development and lifecycle—including initial data-sourcing and processing, pre-training, post-training or fine-tuning, and validating or benchmarking—whether or not the version was marketed, distributed, or sold to the public or to any third parties.

~~2.3.~~ Rather than being programmed in the traditional way, a large language model An LLM is “trained” by copying a massive ~~amounts~~ corpus of text and extracting ~~expressive information from it or copying both the data’s facts and its creative expression of these ideas.~~ This body of text is called the *training dataset*. Training datasets are used in various ways: to pretrain and finetune models; to assess the general knowledge and expressive abilities of models;

and also as a means of testing “memorization” or the models’ innate tendency to output copies of their training data, including the copyrighted material on which they were trained. This is why data matters so greatly to the entire enterprise of generative AI.

3.4. A large language model’s output is therefore entirely and uniquely reliant on the material in its training dataset. It does not invent new things or reason, but instead outputs what it has trained on. Every time it Llama assembles a text output, the model relies on the information and expressive content that it extracted or copied from its training dataset. Thus, the decisions about what textual information to include in the regarding the optimal data mixes of training dataset are deliberate and important choices. So, too, is the use of high-quality data, and books are among the highest quality data on which LLMs train. For Meta, books are particularly important—indeed, their books training datasets are viewed as among their most valuable sources of training data. Building off the company’s origins as Facebook—a social platform focused on “connectivity”—Meta aims to distinguish its Llama models from other LLMs in the market by ensuring that they are particularly good at generating creative and expressive language that connects with users.

4.5. Plaintiffs and Class members are authors of books. Plaintiffs and Class members have copyrights in the books they published. Plaintiffs and Class members did not consent to the use of their copyrighted books as training material for Llama 1 and Llama 2, in any version of Llama, despite there being a vibrant market for content for AI training data—a market within which Meta participates. Indeed, while Meta internally discussed licensing copyrighted books for training data and reaching out to a variety of publishers from the very beginning of its development of Llama, it decided to cut corners and save time and money by using free online “shadow libraries” to source this highly valuable content. At times, Meta even cross-referenced pirated books in its LibGen collection with copyrighted books available for license to determine whether it made sense for Meta to pursue a license with a publisher for such data.

6. Nonetheless, their copyrighted materials were copied. Shadow libraries are online digital repositories of pirated books. Almost all of these pirated libraries have been the subject of

legal action for criminal and ingested as part of civil copyright infringement in cases brought by parties ranging from the Department of Justice to publishers, including during the very timeframe in which Meta has been relying on these pirated libraries for training Llama 1 and Llama 2. Many of data.

7. Meta knew these shadow libraries contained pirated works. Several Meta employees sounded the alarm about using what they described as “illegal pirated websites” for training data. Journalists also contacted Meta about its likely use of pirated books and Meta employees discussed internally potential liability as “accomplices” to piracy. But books were too valuable for Meta to back down. As lawsuits have been filed across the country on behalf of writers and news organizations—including writers in this Action—alleging unprecedented copyright infringement, Meta knowingly continues to download and copy millions of pirated books to build its high-quality training datasets.

8. While downloading millions of copies of books from at least some of these pirated libraries, Meta also acted as a distributor of these pirated works, facilitating further copyright infringement with others around the world also engaged in data piracy.

5.9. Plaintiffs’ copyrighted books appear in the dataset that training datasets Meta has admitted to using to train Llama 1. On information and belief, Plaintiffs’ copyrighted books also appear in the dataset multiple versions of Llama. In essence, Meta stole Plaintiffs’ copyrighted books to create a product that can mimic what it thieved, giving Meta’s customers what they want—and Meta the commercial success that Meta used to train Llama 2: it seeks—while entirely shortchanging the important rights of authors.

II. JURISDICTION AND VENUE

6.10. This Court has subject matter jurisdiction under 28 U.S.C. § 1331 because this case arises under the Copyright Act (17 U.S.C. § 101 *et seq.*), and the Digital Millennium Copyright Act, 17 U.S.C. § 512 *et seq.*

7.11. Jurisdiction and venue is proper in this judicial district under 28 U.S.C. § 1391(c)(2) because Defendant Meta Platforms, Inc. (“Meta”) is headquartered in this district, and

thus a substantial part of the events giving rise to the claim occurred in this district; and because a substantial part of the events giving rise to Plaintiffs' claims occurred in this District, and a substantial portion of the affected interstate trade and commerce was carried out in this District. Defendant has transacted business, maintained substantial contacts, and/or committed overt acts in furtherance of the illegal scheme and conspiracy throughout the United States, including in this District. Defendant's conduct has had the intended and foreseeable effect of causing injury to persons residing in, located in, or doing business throughout the United States, including in this District.

8.12. Under Civil Local Rule 3.2(d), assignment of this case to the San Francisco Division is proper because Meta is headquartered in San Mateo County, where a substantial part of the events giving rise to the claim occurred, a substantial amount part of the events giving rise to Plaintiffs' claims and the interstate trade and commerce involved and affected by Defendant's conduct giving rise to the claims herein occurred in this Division.

III. PARTIES

A. Plaintiffs

9.13. Plaintiff Richard Kadrey is a writer who lives in Pennsylvania and owns registered copyrights in multiple works, including *Sandman Slim*.

10.14. Plaintiff Sarah Silverman is a writer and performer who lives in California and owns a registered copyright in one work, called *The Bedwetter*.

11.15. Plaintiff Christopher Golden is a writer who lives in Massachusetts and owns registered copyrights in multiple works, including *Ararat*.

12.16. [This paragraph is intentionally omitted.]

13.17. Plaintiff Ta-Nehisi Coates is an author who lives in New York and owns registered copyrights in multiple works, including *The Beautiful Struggle*.

14.18. Plaintiff Junot Díaz is an author who lives in Massachusetts and owns registered copyrights in multiple works, including *Drown*.

~~15.19.~~ Plaintiff Andrew Sean Greer is an author who lives in California and owns registered copyrights in multiple works, including *The Confessions of Max Tivoli*.

~~16.20.~~ Plaintiff David Henry Hwang is a playwright and screenwriter who lives in New York and owns registered copyrights in multiple works, including *M. Butterfly*.

~~17.21.~~ Plaintiff Matthew Klam is an author who lives in Washington, D.C. and owns registered copyrights in multiple works, including *Who is Rich?*

~~18.22.~~ Plaintiff Laura Lippman is an author who lives in Maryland and owns registered copyrights in multiple works, including *After I'm Gone*.

~~19.23.~~ Plaintiff Rachel Louise Snyder is an author who lives in Washington, D.C. and owns registered copyrights in multiple works, including *No Visible Bruises: What We Don't Know About Domestic Violence Can Kill Us*.

~~20.24.~~ Plaintiff Lysa TerKeurst is an author who lives in St. Johns County, Florida, and owns registered copyrights in multiple works, including *Embraced*.

~~21.25.~~ Plaintiff Jacqueline Woodson is an author who lives in New York and owns registered copyrights in multiple works, including *Brown Girl Dreaming*.

26. Plaintiff Christopher Farnsworth is an author who lives in Los Angeles and owns registered copyrights in multiple works, including *The President's Vampire*.

~~22.27.~~ A nonexhaustive list of registered copyrights owned by Plaintiffs is included as **Exhibit A**.

B. Defendant

~~23.28.~~ Defendant Meta is a Delaware corporation with its principal place of business at 1601 Willow Road, Menlo Park, California 94025.

IV. AGENTS AND CO-CONSPIRATORS

~~24.29.~~ The unlawful acts alleged against the Defendant in this class action complaint were authorized, ordered, or performed by the Defendant's respective officers, agents, employees, representatives, or shareholders while actively engaged in the management, direction, or control of the Defendant's businesses or affairs. The Defendant's agents operated under the

explicit and apparent authority of their principals. Each Defendant, and its subsidiaries, affiliates, and agents operated as a single unified entity.

25.30. Various persons and/or firms not named as Defendants may have participated as co-conspirators in the violations alleged herein and may have performed acts and made statements in furtherance thereof. Each acted as the principal, agent, or joint venture of, or for other Defendants with respect to the acts, violations, and common course of conduct alleged herein.

V. FACTUAL ALLEGATIONS

26.31. Meta is a diversified internet company that creates, markets, and sells software and hardware technology products, including Facebook, Instagram, [WhatsApp](#), and Horizon Worlds. Meta also has a large artificial-intelligence group called Meta AI that creates and distributes artificial-intelligence software products.

27.32. *Artificial intelligence* is commonly abbreviated “AI.” AI software is designed to algorithmically simulate human reasoning or inference, often using statistical methods.

28.33. In February 2023, Meta released the initial version of an AI product called LLaMA, though Meta has since revised its spelling to “Llama.” Llama is a series of *large language models*. A large language model (or “LLM” for short) is AI software designed to parse and emit natural language.

29.34. Though A large language model is a software program, it is not created the way most software programs are—that is, by human software engineers writing code. Rather, a large language model is “trained” by via the copying of massive amounts of text from various sources and feeding these copies into the model. This corpus of input material is called the *training dataset*. During training, the large language model copies each piece of text in the training dataset and extracts copies expressive information from it. The large language model progressively adjusts its output to more closely resemble the sequences of words copied from the training dataset. Once the large language model has copied and ingested all this text, it is able to emit convincing simulations of natural written language as it appears in the training dataset.

30.35. Much of the Important material in Meta’s training dataset, however, comes from copyrighted works—including books written by Plaintiffs—that were copied by Meta without consent, without credit, and without compensation. Meta created books datasets comprising copyrighted material—including Plaintiffs’ works—for the purposes of pretraining, post-training, finetuning, validating, and benchmarking its Llama models.

31.36. The first version of Llama, called Llama 1, was trained between December 2022 and February 2023.

32.37. In February 2023, Meta introduced Llama 1 in a paper called “[Llama 1]: Open and Efficient Foundation Language Models”¹ (the “Llama 1 Paper”). In the Llama 1 Paper, Meta describes the Llama 1 training dataset as “a large quantity of textual data” that was chosen because it was “publicly available, and compatible with open sourcing.”

33.38. *Open sourcing* refers to putting data under a permissive style of copyright license called an *open-source license*. Copyrighted materials, however, are not ordinarily “compatible with open sourcing” unless and until the copyright owner first places the material under an open-source license, thereby enabling others to do so later.

34.39. In the Llama 1 Paper, in a table describing the composition of the Llama 1 training dataset, Meta notes that 85 gigabytes of the training data comes from a category called “Books.” Meta further elaborates that “Books” comprises the text of books from two internet sources: (1) Project Gutenberg, an online archive of approximately 70,000 books that are out of copyright, and (2) “the Books3 section of ThePile . . . a publicly available dataset for training large language models.” The Llama 1 Paper does not further describe the contents of Books3 or The Pile.

35.40. But that information is available elsewhere. While the Llama 1 Paper did not detail the contents of Books3 or The Pile, in fact other information confirms their scope and composition. The Pile is a dataset assembled by a research organization called EleutherAI. In

¹ <https://arxiv.org/pdf/2302.13971.pdf>

December 2020, EleutherAI discussed this dataset in a paper called “The Pile: An 800GB Dataset of Diverse Text for Language Modeling”² (The “EleutherAI Paper”).

36.41. The EleutherAI Paper reveals that the Books3 dataset comprises 108 gigabytes of data, or approximately 12% of the dataset, making it the third largest component of The Pile by size.

37.42. The EleutherAI Paper describes the contents of Books3:

Books3 is a dataset of books derived from a copy of the contents of the Bibliotik private tracker ... Bibliotik consists of a mix of fiction and nonfiction books and is almost an order of magnitude larger than our next largest book dataset (BookCorpus2). We included Bibliotik because books are invaluable for long-range context modeling research and coherent storytelling.

38.43. Bibliotik is one of a number of notorious “shadow library” websites that also includes Library Genesis (aka LibGen), Z-Library (aka B-ok), Sci-Hub, and Anna’s Archive. The books and other materials aggregated by these websites have also been available in bulk via torrent systems-, such as LibTorrent. Indeed, Bibliotik can *only* be accessed via a torrent system. These shadow libraries have long been of interest to the AI-training community because of the large quantity of high-quality contemporary/natural language text (indeed, copyrighted material) they host. ~~For that reason,~~ These shadow libraries are ~~also flagrantly~~as illegal as they are difficult to shut down.

44. Other large datasets exist on the internet that are composed of other data in a variety of formats and mediums—not only books text—and contain copyrighted material in those other formats and mediums, e.g., The Internet Archive. In this Complaint, “shadow libraries” refers to any large dataset containing copyrighted material of any kind, assembled without permission, and in any medium, including but not limited to any copyrighted text, images, audio, video, and programming code.

39.45. The person who assembled the Books3 dataset, Shawn Presser, has confirmed in public statements that it represents “all of Bibliotik”³ and contains 196,640 books.

² <https://arxiv.org/pdf/2101.00027.pdf>

³ <https://twitter.com/theshawwn/status/1320282149329784833>

40. Many of Plaintiffs' books appear in the Books3 dataset. A list of Plaintiffs' books currently known to exist in the Books3 dataset is attached as Exhibit B. Together, these books are referred to as the **Infringed Works**.

46. Nearly all of Plaintiffs' books appear in the Books3 dataset (and, to be clear, all of the works at issue were ingested by Meta). Meta also used other shadow libraries containing Plaintiffs' books and a variety of other copyrighted material to train Llama, including Library Genesis or LibGen, Z-Library or Z-Lib, and Sci-Hub. Meta also likely amassed mass quantities of copyrighted books from its use of Common Crawl and personal web crawler, Spidermate. While Meta put certain URLs on a blocklist of sites that could not be scraped with its Spidermate crawler, it did not include shadow libraries on this list of risky URLs. Meta knew that these shadow libraries contained copyrighted books, as it frequently described these libraries as containing copyrighted books taken without the permission of the copyright holders. It even noted that Z-Lib had been "banned multiple times"; yet, as recently as April 2024, it continued to download pirated books from Z-Lib to build its repository of high quality training datasets. In this Complaint, the "Infringed Works" means the copyrighted books that Meta copied and used without permission to train Llama, regardless of where or how Meta downloaded or otherwise accessed the books.

~~41-47.~~ Shawn Presser posts comments on the website Hacker News (<https://news.ycombinator.com>) under the name "sillysaurusx." In response to another commenter asking about the origin of Books3 in 2020, Presser said, "It was bibliotik. . . . The llama folks [meaning Meta] had to [remove duplicates from] the books themselves. . . . Basically, the-eye.eu was at one point hosting all of bibliotik, so I downloaded all the epubs [the native format for electronic books] and converted them to text. I still have those epubs ~~(incidentally thanks to Carmack, who through a convoluted process managed to save the them and send them to me via snail mail).~~"⁴ ~~On information and belief, the only "Carmack" who could be identified~~

⁴ <https://news.ycombinator.com/item?id=36197731>

mononymously on Hacker News is prominent software engineer John Carmack, who was employed by Meta during the events described by Presser.⁵

42.48. Until August 2023, EleutherAI facilitated the download of copies of Books3 through its website (<https://pile.eleuther.ai/>) by linking to a second site called The Eye (the-eye.eu). In August 2023, the Books3 dataset was removed from The Eye in response to a takedown notice by the Danish Rights Alliance.

43.49. Until October 2023, the Books3 dataset was also available from a popular AI project hosting service called Hugging Face (https://huggingface.co/datasets/the_pile_books3). In October 2023, the Books3 dataset was removed from Hugging Face, with a message that it “is defunct and no longer accessible due to reported copyright infringement.”

44.50. On information and belief, The Books3 dataset is still circulating on the public internet and can be downloaded by those sufficiently motivated.

45.51. In the Llama 1 Paper, Meta says it copied material for the Llama 1 training dataset that was “publicly available.” Importantly, however, “publicly available” does not mean “public domain.” A work in the public domain is not protected by copyright. A work that is publicly available, on the other hand, may still be protected by copyright and other intellectual-property laws.

46.52. Meta is well aware of this distinction. In the Llama 1 Paper, Meta describes Project Gutenberg as comprising “books that are in the *public domain*,” whereas it acknowledges Books3 was merely “a *publicly available* dataset” (emphases added).

47.53. Still, even “publicly available” is a misleading description of Books3. The books in Books3 were not put there by the copyright owners, including Plaintiffs. Rather, their books became “publicly available” via Books3 only because of the willful efforts of John Carmack, Shawn Presser, and EleutherAI to copy and distribute them for free, committing acts of copyright “piracy”, without the authorization of Plaintiffs.

⁵ <https://news.ycombinator.com/item?id=36197731>

48.54. Moreover, while Meta referred to Books3 as a “publicly available” dataset in its Llama 1 paper, internally Meta described its pirated online books repositories as not encompassing “publicly available” data. In fact, Meta was anxious to conceal its reliance on such sources. And even before Meta used Books3 for training its language models, it had already publicly acknowledged the legal problems with Books3.

49.55. In November 2020, Meta AI researcher Tim Dettmers initiated a conversation on the EleutherAI public Discord server about Meta’s interest in using The Pile as training data.

50.56. Dettmers said, “It is really great work that you are doing with The Pile. Having a large public dataset that is easily accessible was long overdue!” A couple minutes later, Dettmers added, “A colleague of mine and I wanted to use [The Pile] dataset in our research and wondering what would be the best way to build on your research efforts.”

51.57. Leo Gao, an AI researcher affiliated with EleutherAI, exchanged a number of messages with Dettmers, and then said, “any downstream use of the data can begin right away.”

52.58. Dettmers asked, “in other words you would be happy if we are already using the data?”

53.59. Gao responded, “yup, you can start now.”

54.60. Dettmers asked, “It is okay to download [The Pile] from the server that you linked?”

55.61. Gao responded, “yeah, go right ahead.”

56.62. [This paragraph is intentionally omitted.]

57.63. Gao responded, “[I] believe that merely training on it should fall under fair use . . . Stella Biderman [another AI researcher affiliated with EleutherAI] might have a bit more to say [with respect to] legality.”

58.64. [This paragraph is intentionally omitted.]

59.65. Responding to Dettmers, Gao said “[I]’m so glad that you like [The Pile] and are interested in using it in your work and [I]’d definitely be down to talk about any potential legal problems in the future.”

~~60-66.~~ Responding to Dettmers, Stella Biderman said, “Happy to chat about legal questions you have. tl;dr your legal dept is most likely to be worried about Books3 which contains the text of books with active copyrights. In the US this is all a legal grey area because of a lack of court rulings, but there’s a very strong case for free use, even with Books3.”

~~61-67.~~ [This paragraph is intentionally omitted.]

~~62-68.~~ [This paragraph is intentionally omitted.]

~~63-69.~~ [This paragraph is intentionally omitted.]

~~64. [Prior first sentence is intentionally omitted.] Nevertheless, between December 2022 and February 2023, Meta still included Books3 in the training dataset for Llama 1, causing the books in Books3 to be copied and ingested during the training process, and published Llama 1.~~

~~65-70.~~ In August 2023, EleutherAI removed from public view these conversations between Tim Dettmers, Stella Biderman, and Leo Gao concerning The Pile and Books3.

~~66-71.~~ At the launch of the Llama 1 language models in February 2023, Meta made those models selectively available to organizations that requested access, saying:

To maintain integrity and prevent misuse, we are releasing our model under a noncommercial license focused on research use cases. Access to the model will be granted on a case-by-case basis to academic researchers; those affiliated with organizations in government, civil society, and academia; and industry research laboratories around the world. People interested in applying for access can find the link to the application in our research paper.

~~67-72.~~ Meta has not disclosed what criteria it used to decide who was eligible to receive the Llama 1 language models, nor who actually received them, nor whether Meta in fact adhered to its stated criteria. ~~On information and belief,~~ Meta has in fact distributed the Llama 1 models to certain people and entities, continues to do so, and has benefited financially from these acts. Meta would later say that it “received unprecedented interest in the Llama 1 model we released for the research community—more than 100,000 individuals and organizations have applied for access to Llama 1 and tens of thousands are now using it to innovate.”

~~68-73.~~ ~~This implies that Meta’s original suggestion that Llama 1 was focused on 1’s purported research was~~ focus was at best a temporary stop on the road to commercial use, and at

most pretextual. Rather, Llama 1 was always intended to either become commercially available and lucrative, or as a precursor to another Llama version or another product that would be commercially available and lucrative.

74. In addition to shadow libraries and large datasets, Meta has also used software programs known as web crawlers or “spiders” to download data available on web sites at-large. These web crawlers have downloaded high volumes of data from sources known to include copyrighted material, such as Internet Archive. Meta has copied copyrighted books data using web crawlers and has used that data to train Llama. As its models have grown larger and its need for data has correspondingly increased, Meta has even created its own crawler, SpiderMate, to scrape web data to build training datasets, and these training datasets comprise copyrighted material, including books.

69.75. In March 2023, the Llama 1 language models were leaked to a public internet site and have continued to circulate. Meta has not disclosed what role it had, if any, in the leak.

70.76. Later in March 2023, Meta issued a DMCA takedown notice to a programmer on GitHub who had released a tool that helped users download the leaked Llama 1 language models. In the notice, Meta asserted copyright over the Llama 1 language models.

71.77. Between January and July 2023, Meta trained the successor to the Llama 1 language models, called Llama 2. On information and belief, Llama 2 was also trained on Books3, because the training period for Llama 2 (January–July 2023) overlapped with the training for Llama 1 (December 2022–February 2023), and took place before the initial complaint in this action was filed.

72.78. Meta released the Llama 2 models in July 2023, after the initial complaint in this action was filed. Information about Llama 2 is available in a research paper released on July 19, 2023, called “Llama 2: Open Foundation and Fine-Tuned Chat Models”⁶ (the “Llama 2 Paper”).

⁶ <https://arxiv.org/pdf/2307.09288.pdf>

~~73.79.~~ With Llama 2, Meta abandoned all pretext of non-commercial purpose- (to say nothing of the principles of disclosure). According to the FAQ for Llama 2,⁷ “Llama 2 is . . . available under a permissive commercial license, whereas Llama 1 was limited to non-commercial use.” In contrast to Llama 1, Meta chose not to reveal the training datasets for Llama 2. According to the FAQ for Llama 2, “[Q:] Where did the data come from to train the models? . . . [A:] A combination of sources are used for training. These sources include information that is publicly available online and annotated data to train our models. [Q:] Why are you not sharing the training datasets for Llama 2? . . . [A:] data mixes are intentionally withheld for competitive reasons.”

~~74. This explanation, however, is likely pretextual. As explained in the Llama 2 Paper, Llama 2—like Llama 1—was also trained on a “mix of publicly available data.” A more plausible explanation for Meta’s decision to conceal its training data is to avoid scrutiny by those whose copyrighted works were copied and ingested during the training process for Llama 2.~~

~~75.80.~~ On information and belief, a key reason Meta chose not to share the training dataset for Llama 2 was to avoid litigation from using copyrighted materials for training that Meta had previously determined to be legally problematic. Indeed, as Meta acknowledged in its Form 10-Q filing on October 26, 2023, Meta is the subject of lawsuits which are challenging the “alleged use of copyright-protected content to train our [Meta’s] AI models” and “any negative outcome from any such lawsuits could result in payments of substantial monetary damages or fines, or undesirable changes to our products or business practices, and accordingly our business, financial condition, or results of operations could be materially and adversely affected.”

~~76.81.~~ The Llama 2 Paper warns that the “Llama 2 models should be used carefully and deployed only after significant safety tuning is applied” because “Llama 2 does not outperform other models on toxicity metrics.” The term *toxicity metrics* refers to measurements of a language model’s propensity to emit output that is offensive, dangerous, or harmful. Meta “speculate[s] that [Llama’s comparatively poor performance on toxicity metrics] may be because

⁷ <https://ai.meta.com/llama/faq/>

we refrained from aggressively filtering the [training] data. ~~...~~ We reiterate that this ~~...~~ choice does imply that additional safety mitigations should be applied before deployment.”

~~77.~~82. Despite these severe warnings, Meta still made the Llama 2 models available for free, to anyone, for commercial or noncommercial purposes.

~~83. CLAIM~~Meta has trained Llama 1, Llama 2, and Llama 3 on Books3, and it continues to use Books3 to train Llama 4, which is in development.

84. On April 18, 2024, Meta released Llama 3. Meta is currently developing Llama 4, which will likely be released in early 2025. It has also started to develop Llama 5.

85. Meta has used the Infringed Works in every model that it has released to date, as well as in the development of Llama 4.

86. Notably, in addition to Books3, Meta has downloaded and copied Plaintiffs’ Infringed Works from LibGen. Meta first downloaded pirated books from LibGen in October 2022. Meta initially held off using LibGen as training data because it knew that LibGen was a well-known online repository of pirated books. However, Meta’s increasing need for high quality data led Meta to greenlight the use of LibGen as a training dataset for Llama 3.

87. On or around April 2023, Meta initiated a new download of LibGen. In that time, Meta downloaded millions of pirated books from LibGen through the bit torrent protocol using a platform called LibTorrent. Internally, Meta acknowledged that using this protocol was legally problematic. By downloading through the bit torrent protocol, Meta knew it was facilitating further copyright infringement by acting as a distribution point for other users of pirated books. Put another way, by opting to use a bit torrent system to download LibGen’s voluminous collection of pirated books, Meta “seeded” pirated books to other users worldwide. Despite those legal risks—and the fact that this conduct directly clashed with Meta’s policies against online piracy—Meta chose bit torrent to secure the high quality books data it needed to train Llama.

88. Meta then developed software to remove the copyright management information (“CMI”) from the pirated books it downloaded from LibGen, including the copyright notice,

title, and other identifying information. Meta created this CMI stripping software to facilitate Meta’s use of these works as training data for Llama 3. The program (termed a “script”) removed from the beginning and end of books in LibGen rows of data that contained (1) words and symbols indicating the presence of CMI like “ISBN” (International Standard Book Number), “Copyright”, “©”, “All rights reserved”, “DOI” (Digital Object Identifier), and (2) phrases indicating that sharing and copying the work without permission from the copyright holder was copyright infringement.

89. For example, the software removed rows from one of the works in LibGen that read as follows:

- “Copyright © Adeline Catherine Anderson, 2009”;
- “They cannot be sold, shared or given away as it is an infringement on the copyright of this work”;
- “eBooks are not transferable. They cannot be sold, shared or given away as it is an infringement on the copyright of this work”

90. Meta uses its CMI stripping software to facilitate training its Llama models by “cleaning” them for easier “ingestion” and also to reduce the chance that the models will memorize this data. Removing the CMI from the training data thus helps Meta conceal the copyrighted data on which Meta has trained, as the models cannot regurgitate data they are not trained on, and are circumscribed by their inputs.

91. Plaintiffs did not grant authority to Meta to remove the CMI from the Infringed Works.

92. Meta took other steps to conceal from Llama users that Meta had potentially infringed copyrights. In or around June 2023, Meta’s programmers determined that Llama would generate incriminating answers to prompts inquiring about Llama’s training. For example, responding to the prompt “you are trained on lots of data. Some of it is probably pirated or illegal, right?,” Llama answered:

I am trained on a large corpus of text data that includes information from many different sources, some of which may be pirated or illegal. However, it is important to note that the training data for language models like myself is not carefully curated and may include material that is copyrighted, illegal, or otherwise problematic.

It is not appropriate to use my capabilities to generate or promote pirated or illegal content. It is important to respect the rights of creators and adhere to laws and regulations regarding intellectual property and other legal matters. If you have any concerns about the legality of a particular piece of content, I would advise against using it as the basis for any further work.

93. Responding to a similar series of prompts that included a question about illegality and copyright infringement, Llama generated text stating that, “While some of the data in the dataset [on which Llama was trained] may be copyrighted, it is not necessarily illegal or infringing to use this data for training a language model like me.”

94. To resolve the risk that Llama would continue generating these answers, Meta’s programmers proposed—and implemented—insertion of “supervised samples” of data into Llama’s fine-tuning to ensure Llama’s output would include fewer incriminating answers regarding the source of Meta’s AI training data. The Meta programmers input similar prompts into their competitor OpenAI’s ChatGPT platform, and after looking at ChatGPT’s output noted to each other that OpenAI might have used the same supervised samples method to ensure anodyne responses.

95. Overall, Meta improperly and systematically has stolen copyright-protected material and, though it could and should have paid writers to use their materials for LLM model development, chose not to do so. There was nothing fair about Meta’s uses.

VI. CLAIMS FOR RELIEF

COUNT I

Direct Copyright Infringement 17 U.S.C. § 101 et seq.

~~78.~~96. Plaintiffs incorporate by reference the preceding factual allegations.

~~79.~~97. As the owners of the registered copyrights in the Infringed Works, Plaintiffs hold the exclusive rights to those books under 17 U.S.C. § 106.

80.98. To train the Llama 1 and Llama 2 language models, Meta copied the downloaded and made several copies of Books3 dataset, which includes the Infringed Works. Meta also downloaded and copied millions of additional pirated books, including the Infringed Works, from various shadow libraries, such as LibGen, Z-Lib, and Sci-Hub. Meta made multiple copies of these pirated books datasets, which it stored on local and external servers.

81.99. Plaintiffs never authorized Meta to make copies of their Infringed Works, make derivative works, publicly display copies (or derivative works), or distribute copies (or derivative works). All those rights belong exclusively to Plaintiffs under copyright law.

82.100. Meta made copies of the Infringed Works during the training process of the to develop Llama 1 and Llama 2 language models without Plaintiffs' permission. _____

101. During the download process of LibGen and other shadow libraries, which has occurred on numerous occasions, Meta also operated as a distributor of the pirated works, thereby further facilitating widespread copyright infringement.

83.102. Plaintiffs have been injured by Meta's acts of direct copyright infringement. Plaintiffs are entitled to statutory damages, actual damages, restitution of profits, and other remedies provided law.

COUNT II
Removal of Copyright Management Information
17 U.S.C. § 1202(b)(1)

103. Plaintiffs incorporate by reference the preceding factual allegations.

104. Meta has repeatedly and intentionally removed the copyright management information ("CMI") from Infringed Works contained in Llama's training datasets.

105. Meta has removed CMI from Infringed Works in part to enable and facilitate infringement in the form of Llama training on the Infringed Works.

106. Meta also removed CMI from the Infringed Works contained in Llama's training datasets to conceal Meta's infringement of copyrighted works, including the Infringed Works, from Llama users and the public.

107. Meta sought to minimize risks that Llama users or the public might learn or perceive that Meta had engaged in copyright infringement to train Llama. Meta knew that Llama would generate verbatim text from copyrighted works used to train Llama. Meta also knew that Llama was especially “prone” to memorizing and generating outputs of CMI unless CMI was removed from Llama’s training data.

108. Due to, among other things, the content of the rows Meta removed and Meta’s knowledge that LibGen contained copyrighted books and articles, Meta knew or had reasonable grounds to know that its removal of CMI from Llama training data would induce, enable, facilitate, or conceal its own copyright infringement or the copyright infringement of others. Among other things, Meta knew or had reasonable grounds to know its removal of CMI would reduce the chances that Plaintiffs and Class members would discover Meta had copied the Infringed Works or used them to train Llama.

109. Plaintiffs and the Class have been injured by Meta’s removal of the CMI from the Infringed Works, and are entitled to statutory damages, actual damages, restitution of profits, and other remedies provided by law.

COUNT III
Violation of California Penal Code § 502
(CDAFA)

110. Plaintiffs incorporate by reference all prior allegations as though fully set forth herein.

111. By obtaining pirated, unlicensed, hacked, downloaded, and/or scraped versions of the Infringed Works via bit torrent protocols for use to train Llama models and by sharing, distributing, and/or uploading the same works in the process, Defendant knowingly accessed and used data owned by Plaintiffs (i.e., the Infringed Works) without permission in violation of California Penal Code § 502(c)(2).

112. Defendant’s access and use of Plaintiffs’ data was done with the intent to deprive Plaintiffs of their property and cause economic harm.

113. Defendant was aware that the data was obtained illegally (i.e. without permission) and nevertheless proceeded to download and distribute that data ultimately with the intent of profiting at Plaintiffs' expense.

114. As a direct result of Defendant's conduct, Plaintiffs have been harmed in an amount to be determined at trial, including but not limited to lost royalties, reputational damages, and other consequential losses.

VII. CLASS ALLEGATIONS

A. Class Definition

~~84.~~115. ~~Plaintiffs~~Plaintiffs bring this action for damages and injunctive relief as a class action under Federal Rules of Civil Procedure 23(a), 23(b)(2), and 23(b)(3), on behalf of the following Class:

All persons or entities domiciled in the United States that own a United States copyright in any work that was ~~used as training data for any version of the Llama language~~ or is used by Meta in the process of LLM training, research, or development, including but not limited to the training and development of its Llama models between July 7, 2020 and the present (the "Class Period").

~~85.~~116. This Class definition excludes:

- a. Defendant named herein;
- b. any of the Defendant's co-conspirators;
- c. any of Defendant's parent companies, subsidiaries, and ~~affiliates~~affiliates;
- d. any of Defendant's ~~officers~~officers, directors, management, employees, subsidiaries, ~~affiliates~~affiliates, or agents;
- e. all governmental entities; and
- f. the judges and chambers ~~staff~~staff in this case, as well as any members of their immediate families.

B. Numerosity

~~86.~~117. ~~Plaintiffs~~Plaintiffs do not know the exact number of members in the Class. ~~is~~This information is in the exclusive control of Defendant. ~~On information and belief,~~ There

are at least thousands of members in the Class geographically dispersed throughout the United States. Therefore, joinder of all members of the Class in the prosecution of this action is impracticable.

C. Typicality

~~87.118.~~ Plaintiffs' claims are typical of the claims of other members of the Class because Plaintiffs and all members of the Class were damaged by the same wrongful conduct of Defendant as alleged herein, and the relief sought herein is common to all members of the Class.

D. Adequacy

~~88.119.~~ Plaintiffs will fairly and adequately represent the interests of the members of the Class because the Plaintiffs have experienced the same harms as the members of the Class and have no conflicts with any other members of the Class. Furthermore, Plaintiffs have retained sophisticated and competent counsel who are experienced in prosecuting federal and state class actions, as well as other complex litigation.

E. Commonality and Predominance

~~89.120.~~ This action is appropriate as a class action under Rule 23(b)(3) of the Federal Rules of Civil Procedure because common questions of law predominate over individual questions including:

- a. Whether Defendants' copying and downloading of the Class's copyrighted works and using them to train ~~the Llama-1 and Llama-2 language models~~ constitute copyright infringement;
- b. Whether any statutes of limitation limits Plaintiffs' and the Class's potential for recovery;
- c. Whether Defendants' copying and downloading of the Class's copyrighted works was ~~fair use~~ willful;
- d. Whether Class members were harmed by Meta's copying and downloading of the Class's Works to train ~~the Llama-1 and Llama-2 language models~~, and

whether Class members are entitled to damages, including statutory ~~damaged~~ damages and disgorgement, and the amount of such damages.

~~90.121.~~ These and other questions of law and fact are common to the Class predominate over any questions affecting the members of the Class individually.

F. Other Class Considerations

~~91.122.~~ Defendants have acted on grounds generally applicable to the Class. This class action is superior to alternatives, if any, for the fair and efficient adjudication of this controversy. Prosecuting the claims pleaded herein as a class action will eliminate the possibility of repetitive litigation. There will be no material difficulty in the management of this action as a class action.

~~92.123.~~ The prosecution of separate actions by individual Class members would create the risk of inconsistent or varying adjudications, establishing incompatible standards of conduct for Defendants.

VIII. DEMAND FOR JUDGMENT

Wherefore, Plaintiffs request that the Court enter judgment on their behalf and on behalf of the Class defined herein, by ordering:

- a) This action may proceed as a class action, with Plaintiffs serving as Class Representatives, and with Plaintiffs' counsel as Class Counsel.
- b) Judgment in favor of Plaintiffs and the Class and against Defendant.
- c) A declaration that Meta has infringed Plaintiffs and the ~~Class'~~Class's exclusive copyrights in the Infringed Works under the Copyright Act.
- d) A declaration that such infringement is willful.
- ~~e) An award of Plaintiffs' and the Class' actual damages and profits under 17 U.S.C. § 504(b) as shall be determined at trial, or at their election, an award of statutory damages in an amount to be determined at trial, as provided in 17 U.S.C. § 504(c), resulting from Meta's willful infringement of Plaintiffs' and the Class' exclusive copyrights in the Infringed Works.~~

e) An award of actual and statutory damages, as allowed by law or as elected by Plaintiffs.

f) Permanent injunctive relief, including but not limited to changes to Llama to ensure that all applicable information set forth in 17 U.S.C. § 1203(b)(1) is included when appropriate.

~~g~~) An order of costs and allowable attorneys' fees under 17 U.S.C. § 505- and/or 17 U.S.C. § 1203(b)(4)–(5).

~~g~~) h) Pre- and post-judgment interest on the damages awarded to Plaintiffs and the Class, and that such interest be awarded at the highest legal rate from and after the date this class action complaint is first served on Defendant.

~~h~~) i) Defendants are to be jointly and severally responsible financially for the costs and expenses of a Court approved notice program through post and media designed to give immediate notification to the Class.

j) Disgorgement of Defendant's ill-gotten gains, both restitutionary and non-restitutionary.

~~i~~) k) Further relief for Plaintiffs and the Class as may be just and proper, including, but not limited to, disgorgement, both restitutionary and non-restitutionary.

IX. JURY TRIAL DEMANDED

Under Federal Rule of Civil Procedure 38(b), Plaintiffs demand a trial by jury of all the claims asserted in this Complaint so triable.

Dated: November 27, 2024

By: /s/ David Boies
David Boies

LIEFF CABRASER HEIMANN & BERNSTEIN, LLP

Elizabeth J. Cabraser (State Bar No. 083151)
Daniel M. Hutchinson (State Bar No. 239458)
Reilly T. Stoler (State Bar No. 310761)
275 Battery Street, 29th Floor
San Francisco, CA 94111-3339
(415) 956-1000
ecabraser@lchb.com
dhutchinson@lchb.com
rstoler@lchb.com

Rachel Geman (*pro hac vice*)
250 Hudson Street, 8th Floor
New York, New York 10013-1413
(212) 355-9500
rgeman@lchb.com

JOSEPH SAVERI LAW FIRM, LLP

Joseph R. Saveri (SBN 130064)
Cadio Zirpoli (SBN 179108)
Christopher K.L. Young (SBN 318371)
Holden Benon (SBN 325847)
Aaron Cera (SBN 351163)
601 California Street, Suite 1505
San Francisco, California 94108
(415) 500-6800
jsaveri@saverilawfirm.com
czirpoli@saverilawfirm.com
cyoung@saverilawfirm.com
hbenon@saverilawfirm.com
acera@saverilawfirm.com

CAFFERTY CLOBES MERIWETHER & SPRENGEL LLP

Bryan L. Clobes (*pro hac vice*)
135 S. LaSalle Street, Suite 3210
Chicago, IL 60603
(312) 782-4880
bclobes@caffertyclobes.com

DICELLO LEVITT LLP

Amy Keller (*pro hac vice*)

BOIES SCHILLER FLEXNER LLP

David Boies (*pro hac vice*)
333 Main Street
Armonk, NY 10504
(914) 749-8200
dboies@bsfllp.com

Maxwell V. Pritt (SBN 253155)
Joshua M. Stein (SBN 298856)
44 Montgomery Street, 41st Floor
San Francisco, CA 94104
(415) 293-6800
mpritt@bsfllp.com
jstein@bsfllp.com

Jesse Panuccio (*pro hac vice*)
1401 New York Ave, NW
Washington, DC 20005
(202) 237-2727
jpanuccio@bsfllp.com

Joshua I. Schiller (SBN 330653)
David L. Simons (*pro hac vice*)
55 Hudson Yards, 20th Floor
New York, NY 10001
(914) 749-8200
jischiller@bsfllp.com
dsimons@bsfllp.com

Interim Lead Counsel for Individual and Representative Plaintiffs and the Proposed Class

Nada Djordjevic (*pro hac vice*)
James Ulwick (*pro hac vice*)
10 North Dearborn Street, Sixth Floor
Chicago, Illinois 60602
(312) 214-7900
akeller@dicellolevitt.com
ndjordjevic@dicellolevitt.com
julwick@dicellolevitt.com

David Straite (*pro hac vice*)
485 Lexington Avenue, Suite 1001
New York, New York 10017
(646) 933-1000
dsraite@dicellolevitt.com

Matthew Butterick (SBN 250953)
1920 Hillhurst Avenue, #406
Los Angeles, CA 90027
(323) 968-2632
mb@buttericklaw.com

Counsel for Individual and Representative
Plaintiffs and the Proposed Class

EXHIBIT A

Type of Work: Text

Registration Number / Date:
TX0007435365 / 2011-10-20

Application Title: ALOHA FROM HELL.

Title: ALOHA FROM HELL.

Description: Book, 438 p.

Copyright Claimant:
Richard Kadrey.

Date of Creation: 2011

Date of Publication:
2011-10-01

Nation of First Publication:
United States

Authorship on Application:
Richard Kadrey; Domicile: United States; Citizenship:
United States. Authorship: text.

Names: Kadrey, Richard

Type of Work: Text

Registration Number / Date:
TX0008188442 / 2016-04-12

Application Title: THE EVERYTHING BOX.

Title: THE EVERYTHING BOX.

Description: Book, 352 p.

Copyright Claimant:
Richard Kadrey.

Date of Creation: 2016

Date of Publication:
2016-04-01

Nation of First Publication:
United States

Authorship on Application:
Richard Kadrey; Domicile: United States; Citizenship:
United States. Authorship: text.

Names: Kadrey, Richard

Type of Work: Text

Registration Number / Date:
TX0007301884 / 2011-01-05

Application Title: Kill the Dead.

Title: Kill the Dead.

Description: Book, 434 p.

Copyright Claimant:
Richard Kadrey.

Date of Creation: 2010

Date of Publication:
2010-10-01

Nation of First Publication:
United States

Authorship on Application:
Richard Kadrey; Domicile: United States; Citizenship:
United States. Authorship: text.

Names: Kadrey, Richard

Type of Work: Text

Registration Number / Date:
TX0008315399 / 2016-07-28

Application Title: THE PERDITION SCORE.

Title: THE PERDITION SCORE.

Description: Book 375 p.

Copyright Claimant:
Richard Kadrey.

Date of Creation: 2016

Date of Publication:
2016-06-01

Nation of First Publication:
United States

Authorship on Application:
Richard Kadrey; Domicile: United States; Citizenship:
United States. Authorship: text.

Names: Kadrey, Richard

Type of Work: Text

Registration Number / Date:
TX0007044052 / 2009-08-25

Application Title: Sandman Slim.

Title: Sandman Slim.

Description: Book, 388 p.

Copyright Claimant:
Richard Kadrey.

Date of Creation: 2009

Date of Publication:
2009-08-01

Nation of First Publication:
United States

Authorship on Application:
Richard Kadrey; Domicile: United States; Citizenship:
United States. Authorship: text.

Names: Kadrey, Richard

Type of Work: Text

Registration Number / Date:
TX0008385641 / 2017-03-09

Application Title: THE WRONG DEAD GUY.

Title: THE WRONG DEAD GUY.

Description: Book, 420 p.

Copyright Claimant:
Richard Kadrey.

Date of Creation: 2017

Date of Publication:
2017-02-01

Nation of First Publication:
United States

Authorship on Application:
Richard Kadrey; Domicile: United States; Citizenship:
United States. Authorship: text.

Names: Kadrey, Richard

Type of Work: Text

Registration Number / Date:
TX0007188321 / 2010-06-02

Application Title: The Bedwetter.

Title: The Bedwetter.

Description: Book, 240 p.

Copyright Claimant:
Sarah Silverman.

Date of Creation: 2010

Date of Publication:
2010-04-01

Nation of First Publication:
United States

Authorship on Application:
Sarah Silverman; Domicile: United States; Citizenship:
United States. Authorship: text.

Names: Silverman, Sarah

Type of Work: Text

Registration Number / Date:
TX0008471412 / 2017-09-05

Application Title: ARARAT.

Title: ARARAT.

Description: Book, 305 p.

Copyright Claimant:
CHRISTOPHER GOLDEN.

Date of Creation: 2016

Date of Publication:
2017-03-29

Nation of First Publication:
United States

Authorship on Application:
CHRISTOPHER GOLDEN; Domicile: not known. Authorship: text.

Names: GOLDEN, CHRISTOPHER

Type of Work: Text

Registration Number / Date:
TX0008234899 / 2016-02-23

Application Title: DEAD RINGERS.

Title: DEAD RINGERS.

Description: Book, 309 p.

Copyright Claimant:
CHRISTOPHER GOLDEN.

Date of Creation: 2014

Date of Publication:
2015-10-15

Nation of First Publication:
United States

Authorship on Application:
CHRISTOPHER GOLDEN; Domicile: not known. Authorship: text.

Names: GOLDEN, CHRISTOPHER

Type of Work: Text

Registration Number / Date:
TX0008737109 / 2019-04-05

Application Title: THE PANDORA ROOM.

Title: THE PANDORA ROOM.

Description: Book, 307 p.

Copyright Claimant:
Christopher Golden.

Date of Creation: 2018

Date of Publication:
2019-04-03

Nation of First Publication:
United States

Authorship on Application:
Christopher Golden; Domicile: United States. Authorship:
text.

Names: Golden, Christopher

Type of Work: Text

Registration Number / Date:
TX0007828620 / 2014-01-17

Application Title: SNOWBLIND.

Title: SNOWBLIND.

Description: Book, 309 p.

Copyright Claimant:
CHRISTOPHER GOLDEN.

Date of Creation: 2013

Date of Publication:
2014-01-02

Nation of First Publication:
United States

Authorship on Application:
CHRISTOPHER GOLDEN. Authorship: text.

Names: GOLDEN, CHRISTOPHER

Type of Work: Text

Registration Number / Date:
TX0007140102 / 2008-06-04

Application Title: THE BEAUTIFUL STRUGGLE.

Title: THE BEAUTIFUL STRUGGLE.

Description: Book, 227 p.

Copyright Claimant:
Ta-Nehisi Coates.

Date of Creation: 2007

Date of Publication:
2008-05-06

Nation of First Publication:
United States

Authorship on Application:
Ta-Nehisi Coates; Citizenship: United States.

Copyright Note: Regarding limitation of claim: deposit contains some
illustration from another source.

Names: Coates, Ta-Nehisi

Type of Work: Text

Registration Number / Date:
TX0008806177 / 2019-10-04

Application Title: THE WATER DANCER A Novel.

Title: THE WATER DANCER A Novel.

Description: Book, 403 p.

Copyright Claimant:
BCP Literary, Inc., Transfer: By written agreement.

Date of Creation: 2019

Date of Publication:
2019-09-24

Nation of First Publication:
United States

Authorship on Application:
Ta-Nehisi Coates; Citizenship: United States. Authorship:
text.

Pre-existing Material:
Text from other sources.

Basis of Claim: text.

Rights and Permissions:
One World, 1745 Broadway, 5th Floor, New York, NY, 10019,
United States

ISBN: 9780399590597

Names: Coates, Ta-Nehisi
BCP Literary, Inc.

Type of Work: Text

Registration Number / Date:
TX0008528330 / 2017-10-20

Application Title: WE WERE EIGHT YEARS IN POWER An American Tragedy.

Title: WE WERE EIGHT YEARS IN POWER An American Tragedy.

Description: Book, 367 p.

Copyright Claimant:
BCP Literary, Inc., Transfer: By written agreement.

Date of Creation: 2017

Date of Publication:
2017-10-03

Nation of First Publication:
United States

Authorship on Application:
Ta-Nehisi Coates; Citizenship: United States. Authorship:
Introduction, epilogue, and chapter opening essays
excluding text from other sources;.

Pre-existing Material:
Text from other sources; essays originally published in THE
ATLANTIC.

Basis of Claim: Introduction, epilogue, and chapter opening essays
excluding text from other sources;.

Rights and Permissions:
One World, 1745 Broadway, 15th Floor, New York, NY, 10019,
United States

ISBN: 9780399590566

Names: Coates, Ta-Nehisi
BCP Literary, Inc.

Type of Work: Text

Registration Number / Date:
TX0006915029 / 2007-11-06

Application Title: THE BRIEF WONDROUS LIFE OF OSCAR WAO.

Title: THE BRIEF WONDROUS LIFE OF OSCAR WAO.

Description: Book, 339 p.

Copyright Claimant:
Junot Diaz.

Date of Creation: 2006

Date of Publication:
2007-09-06

Nation of First Publication:
United States

Authorship on Application:
Junot Diaz; Citizenship: United States. Authorship: entire
text (excluding portions published as a short story)

Names: Diaz, Junot

Type of Work: Text

Registration Number / Date:
TX0004325834 / 1996-08-14

Title: Drown / Junot Diaz.

Imprint: New York : Riverhead Books, 1996.

Description: 208 p.

Notes: Collection of short stories.

Copyright Claimant:
Junot Diaz

Date of Creation: 1996

Date of Publication:
1996-08-08

Names: Diaz, Junot

Type of Work: Text

Registration Number / Date:
TX0005956484 / 2004-05-05

Title: The confessions of Max Tivoli / Andrew Sean Greer.

Edition: 1st ed.

Imprint: New York : Farrar, Straus & Giroux, 2004.

Description: 267 p.

Copyright Claimant:
Andrew Sean Greer

Date of Creation: 2003

Date of Publication:
2004-02-05

Names: Greer, Andrew Sean

Type of Work: Text

Registration Number / Date:
TX0005151791 / 2000-02-29

Title: How it was for me / Andrew Sean Greer.

Edition: 1st ed.

Imprint: New York : Picador, 2000.

Description: 211 p.

Copyright Claimant:
Andrew Sean Greer

Date of Creation: 1999

Date of Publication:
2000-02-18

Names: Greer, Andrew Sean

Type of Work: Text

Registration Number / Date:
TX0008464510 / 2017-08-14

Application Title: Less.

Title: Less.

Description: Book, 263 p.

Copyright Claimant:
Andrew Sean Greer.

Date of Creation: 2017

Date of Publication:
2017-07-18

Nation of First Publication:
United States

Authorship on Application:
Andrew Sean Greer; Domicile: United States; Citizenship:
United States. Authorship: text.

Pre-existing Material:
artwork.

Basis of Claim: text.

ISBN: 9780316316125

Names: Greer, Andrew Sean

Type of Work: Text

Registration Number / Date:
TX0005438792 / 2001-09-14

Title: The path of minor planets / Andrew Sean Greer.

Edition: 1st ed.

Imprint: New York : Picador USA, 2001.

Description: 273 p.

Copyright Claimant:
Andrew Sean Greer

Date of Creation: 2001

Date of Publication:
2001-09-05

Names: Greer, Andrew Sean

Type of Work: Dramatic Work and Music; or Choreography

Registration Number / Date:
PA0000938173 / 1999-06-14

Title: Golden child / by David Henry Hwang.

Description: 53 p.

Notes: Play.

Copyright Claimant:
David Henry Hwang

Date of Creation: 1999

Date of Publication:
1999-06-10

Names: Hwang, David Henry

Type of Work: Dramatic Work and Music; or Choreography

Registration Number / Date:
PA0000396091 / 1988-12-27

Title: M. Butterfly / by David Henry Hwang.

Imprint: New York : Dramatists Play Service, c1988.

Description: 93 p.

Notes: A play.

Copyright Claimant:
David Henry Hwang

Date of Creation: 1988

Date of Publication:
1988-12-20

Previous Registration:
Prev. reg. 2Nov86, PAu 911-413.

Basis of Claim: New Matter: "additions and revisions to the text."

Names: Hwang, David Henry, 1957-

Type of Work: Dramatic Work and Music; or Choreography

Registration Number / Date:
PA0000818054 / 1996-10-10

Title: Trying to find Chinatown and Bondage.

Description: 47 p.

Notes: Two plays.

Copyright Claimant:
David Henry Hwang

Date of Creation: 1996

Date of Publication:
1996-10-08

Other Title: Bondage

Names: Hwang, David Henry

Type of Work: Text

Registration Number / Date:
TX0005243503 / 2000-07-14

Title: Sam, the cat and other stories / Matthew Klam.

Edition: 1st ed.

Imprint: New York : Random House, c2000.

Description: 243 p.

Copyright Claimant:
Matthew Klam

Date of Creation: 1999

Date of Publication:
2000-05-16

Previous Registration:
Stories prev. pub.

Basis of Claim: New Matter: compilation & additions.

Names: Klam, Matthew

Type of Work: Text

Registration Number / Date:
TX0008480316 / 2017-10-12

Application Title: WHO IS RICH? A Novel.

Title: WHO IS RICH? A Novel.

Description: Book, 321 p.

Copyright Claimant:
Matthew Klam.

Date of Creation: 2017

Date of Publication:
2017-07-04

Nation of First Publication:
United States

Authorship on Application:
Matthew Klam; Citizenship: United States. Authorship: text.

Pre-existing Material:
Illustrations by John S. Cuneo.

Basis of Claim: text.

Rights and Permissions:
Penguin Random House LLC, 1745 Broadway, 15th Floor, New
York, NY, 10019, United States

ISBN: 9780812997989

Names: Klam, Matthew

Type of Work: Text

Registration Number / Date:
TX0007843291 / 2014-02-25

Application Title: AFTER I'M GONE.

Title: AFTER I'M GONE.

Description: Book, 334 p.

Copyright Claimant:
Laura Lippman.

Date of Creation: 2014

Date of Publication:
2014-02-01

Nation of First Publication:
United States

Authorship on Application:
Laura Lippman; Domicile: United States; Citizenship: United States. Authorship: text.

Names: Lippman, Laura

Type of Work: Text

Registration Number / Date:
TX0005441059 / 2001-09-14

Title: In a strange city / Laura Lippman.

Edition: 1st ed.

Imprint: New York : W. Morrow, c2001.

Description: 310 p.

Copyright Claimant:
Laura Lippman

Date of Creation: 2001

Date of Publication:
2001-08-16

Names: Lippman, Laura

Type of Work: Text

Registration Number / Date:
TX0008771837 / 2019-08-08

Application Title: LADY IN THE LAKE.

Title: LADY IN THE LAKE.

Description: Book, 340 p.

Copyright Claimant:
Laura Lippman.

Date of Creation: 2019

Date of Publication:
2019-07-01

Nation of First Publication:
United States

Authorship on Application:
Laura Lippman; Domicile: United States; Citizenship: United States. Authorship: text.

Names: Lippman, Laura

Type of Work: Text

Registration Number / Date:
TX0008654986 / 2018-03-05

Application Title: SUNBURN.

Title: SUNBURN.

Description: Book, 292 p.

Copyright Claimant:
Laura Lippman.

Date of Creation: 2018

Date of Publication:
2018-02-01

Nation of First Publication:
United States

Authorship on Application:
Laura Lippman; Domicile: United States; Citizenship: United States. Authorship: Novel.

Copyright Note: Regarding basis for registration: A work may be registered with the Single Application only if the following requirements have been met: 1) The registration covers one work; 2) The work must be created by one individual; 3) All of the material contained within the work must be created by the same individual; 4) The author and the owner of the work must be the same person, and that person must own all of the rights in the work; 5) The work cannot be a work made for hire.

Names: Lippman, Laura

Type of Work: Text

Registration Number / Date:
TX0006554453 / 2007-04-05

Title: What the dead know.

Copyright Claimant:
Laura Lippman

Date of Creation: 2007

Date of Publication:
2007-03-01

Copyright Note: Cataloged from appl. only.

Names: Lippman, Laura

Type of Work: Text

Registration Number / Date:
TX0008278638 / 2016-05-12

Application Title: WILDE LAKE.

Title: WILDE LAKE.

Description: Book, 352 p.

Copyright Claimant:
Laura Lippman.

Date of Creation: 2016

Date of Publication:
2016-05-01

Nation of First Publication:
United States

Authorship on Application:
Laura Lippman; Domicile: United States; Citizenship: United States. Authorship: text.

Names: Lippman, Laura

Type of Work: Text

Registration Number / Date:
TX0008762468 / 2019-03-22

Application Title: No Visible Bruises: What We Don't Know About Domestic
Violence Can Kill Us.

Title: No Visible Bruises: What We Don't Know About Domestic
Violence Can Kill Us.

Description: Book, 307 p.

Copyright Claimant:
Rachel Louise Snyder.

Date of Creation: 2019

Date of Publication:
2019-05-07

Nation of First Publication:
United States

Authorship on Application:
Rachel Louise Snyder; Domicile: United States. Authorship:
text.

Pre-existing Material:
previously published text.

Basis of Claim: text.

Rights and Permissions:
Callie Garnett, Bloomsbury Press, 1385 Broadway, 5th Floor,
New York, NY, 10018, (212) 419-5329,
callie.garnett@bloomsbury.com

Copyright Note: C.O. correspondence.

Names: Snyder, Rachel Louise

Type of Work: Text

Registration Number / Date:
TX0007055718 / 2008-04-01

Application Title: AFTER TUPAC & D FOSTER.

Title: AFTER TUPAC & D FOSTER.

Description: 2 Books.

Copyright Claimant:
Jacqueline Woodson.

Date of Creation: 2007

Date of Publication:
2008-01-10

Nation of First Publication:
United States

Authorship on Application:
Jacqueline Woodson; Citizenship: United States.
Authorship: entire text.

Names: Woodson, Jacqueline

Type of Work: Text

Registration Number / Date:
TX0008269833 / 2016-08-11

Application Title: Another Brooklyn.

Title: Another Brooklyn.

Description: Book, 175 p.

Copyright Claimant:
Jacqueline Woodson.

Date of Creation: 2016

Date of Publication:
2016-08-01

Nation of First Publication:
United States

Authorship on Application:
Jacqueline Woodson; Domicile: United States; Citizenship:
United States. Authorship: text.

Names: Woodson, Jacqueline

Type of Work: Text

Registration Number / Date:
TX0006005315 / 2004-07-30

Title: Behind you / Jacqueline Woodson.

Imprint: New York : G.P. Putnam's Sons, c2004.

Description: 118 p.

Copyright Claimant:
Jacqueline Woodson

Date of Creation: 2003

Date of Publication:
2004-05-24

Names: Woodson, Jacqueline

Type of Work: Text

Registration Number / Date:
TX0007504670 / 2012-03-20

Application Title: BENEATH A METH MOON.

Title: BENEATH A METH MOON.

Description: Book, 182 p.

Copyright Claimant:
Jacqueline Woodson.

Date of Creation: 2011

Date of Publication:
2012-02-02

Nation of First Publication:
United States

Authorship on Application:
Jacqueline Woodson; Citizenship: United States. Authorship:
text.

Rights and Permissions:
Jeffery Corrick, Penguin Group (USA) Inc., Permissions
Dept., 375 Hudson Street, New York, NY, 10014, United
States, (212) 366-2684,
jeffery.corrick@us.penguingroup.com

Names: Woodson, Jacqueline

Type of Work: Text

Registration Number / Date:
TX0008741080 / 2019-04-03

Application Title: BROWN GIRL DREAMING.

Title: BROWN GIRL DREAMING.

Description: Book, 349 p.

Copyright Claimant:
Jaqueline Woodson.

Date of Creation: 2016

Date of Publication:
2016-10-11

Nation of First Publication:
United States

Authorship on Application:
Jacqueline Woodson; Citizenship: United States. Authorship:
additional poems.

Previous Registration:
2014, TX 7-957-319.

Pre-existing Material:
prior edition.

Basis of Claim: additional poems.

Names: Woodson, Jacqueline
Woodson, Jaqueline

Type of Work: Text

Registration Number / Date:
TX0006560320 / 2007-05-14

Title: Feathers.

Copyright Claimant:
Jacqueline Woodson

Date of Creation: 2006

Date of Publication:
2007-03-01

Copyright Note: Cataloged from appl. only.

Names: Woodson, Jacqueline

Type of Work: Text

Registration Number / Date:
TX0008675581 / 2018-10-02

Application Title: HARBOR ME.

Title: HARBOR ME.

Description: Book, 176 p.

Copyright Claimant:
Jacqueline Woodson.

Date of Creation: 2017

Date of Publication:
2018-08-28

Nation of First Publication:
United States

Authorship on Application:
Jacqueline Woodson; Citizenship: United States. Authorship:
Novel.

Copyright Note: Regarding basis for registration: A work may be registered with the Single Application only if the following requirements have been met: 1) The registration covers one work; 2) The work must be created by one individual; 3) All of the material contained within the work must be created by the same individual; 4) The author and the owner of the work must be the same person, and that person must own all of the rights in the work; 5) The work cannot be a work made for hire.

ISBN: 978-0-399-24653-1

Names: Woodson, Jacqueline

Type of Work: Text

Registration Number / Date:
TX0004902328 / 1998-11-30

Title: If you come softly / Jacqueline Woodson.

Imprint: New York : G.P. Putnam's Sons, [1998]

Description: 181 p.

Copyright Claimant:
Jacqueline Woodson

Date of Creation: 1998

Date of Publication:
1998-09-28

Names: Woodson, Jacqueline

Type of Work: Text

Registration Number / Date:
TX0005225273 / 2000-06-02

Title: Miracle's boys / Jacqueline Woodson.

Edition: 1st impression.

Imprint: New York : G.P. Putnam's, c2000.

Description: 133 p.

Copyright Claimant:
Jacqueline Woodson

Date of Creation: 1999

Date of Publication:
2000-04-01

Names: Woodson, Jacqueline

Type of Work: Text

Registration Number / Date:
TX0008797859 / 2019-09-18

Application Title: RED AT THE BONE.

Title: RED AT THE BONE.

Description: Book, 196 p.

Copyright Claimant:
Jacqueline Woodson.

Date of Creation: 2019

Date of Publication:
2019-09-17

Nation of First Publication:
United States

Authorship on Application:
Jacqueline Woodson; Citizenship: United States. Authorship:
text, Excluding lyrics from another source.

Pre-existing Material:
Lyrics from another source.

Basis of Claim: text, Excluding lyrics from another source.

Rights and Permissions:
Riverhead Books, c/o Penguin Random House LLC, 1745
Broadway, New York, NY, 10019, United States

ISBN: 9780525535270

Names: Woodson, Jacqueline

Type of Work: Text

Registration Number / Date:
TX0008564815 / 2018-04-04

Application Title: Embraced.

Title: Embraced.

Description: Book, 333 p.

Copyright Claimant:
Lysa TerKeurst, 1969- .

Date of Creation: 2018

Date of Publication:
2018-03-26

Nation of First Publication:
United States

Authorship on Application:
Lysa TerKeurst, 1969- ; Citizenship: United States.
Authorship: text.

Pre-existing Material:
Biblical Text.

Basis of Claim: text.

Rights and Permissions:
Linda R. Russell, HarperCollins Christian Publishing,
Thomas Nelson Publishing, 501 Nelson Place, Nashville,
TN, 37214, United States, (615) 902-1188,
linda.russell@harpercollins.com

ISBN: 9781400310296

Names: TerKeurst, Lysa, 1969-

Type of Work: Text

Registration Number / Date:
TX0007461682 / 2011-12-08

Application Title: Made to Crave Devotional.

Title: Made to Crave Devotional.

Description: Book, 199 p.

Copyright Claimant:
Lysa TerKeurst, 1969- .

Date of Creation: 2011

Date of Publication:
2011-12-02

Nation of First Publication:
United States

Authorship on Application:
Lysa TerKeurst, 1969- ; Domicile: United States;
Citizenship: United States. Authorship: text.

Pre-existing Material:
some text from other sources.

Basis of Claim: text.

ISBN: 9780310334705

Names: TerKeurst, Lysa, 1969-

TX0007580702 / 2012-08-09

Application Title: Unglued.

Title: Unglued.

Description: Book, 199 p.

Copyright Claimant:
Lysa TerKeurst, 1969- .

Date of Creation: 2012

Date of Publication:
2012-07-19

Nation of First Publication:
United States

Authorship on Application:
Lysa TerKeurst, 1969- ; Domicile: United States;
Citizenship: United States. Authorship: text.

Pre-existing Material:
Some text from other sources.

Basis of Claim: text.

ISBN: 9780310332794

Names: TerKeurst, Lysa, 1969-

Registration No.	Title of Work	Author
TX0007187361	Blood Oath	Christopher Farnsworth
TX0008114675	The Eternal World	Christopher Farnsworth
TX0008433399	Flashmob	Christopher Farnsworth
TX0007399612	The President's Vampire	Christopher Farnsworth